

Lexical Awareness and Development through Data Driven Learning: Attitudes and Beliefs of EFL Learners

Asuman Aşık¹, Arzu Şarlanoğlu Vural¹, Kadriye Dilek Akpınar¹

¹Department of English Language Teaching, Gazi University, Turkey

Correspondence: Asuman Aşık, Department of English Language Teaching, Gazi University, Beşevler, 06500, Ankara, Turkey.

Received: November 30, 2015 Accepted: December 14, 2015 Online Published: December 23, 2015

doi:10.11114/jets.v4i3.1223

URL: <http://dx.doi.org/10.11114/jets.v4i3.1223>

Abstract

Data-driven learning (DDL) has become an innovative approach developed from corpus linguistics. It plays a significant role in the progression of foreign language pedagogy, since it offers learners plentiful authentic corpora examples that make them analyze language rules with the help of online corpora and concordancers. The present study attempts to reveal Turkish EFL learners' attitudes and beliefs about DDL instruction in terms of lexical awareness and development. A mixed method including a questionnaire and a focus group interview was used to gather data. The statistical analysis of the results indicated positive attitudes towards DDL instruction in terms of raising the depth of lexical awareness particularly for synonyms and collocations. However, participants reported that their awareness towards word frequency, idioms and vocabulary learning strategies did not increase satisfactorily. Negative attitudes were also observed due to technical problems of the software and time-consuming nature of doing tasks through corpus.

Keywords: corpus, Data Driven Learning, lexical competence, language learning and teaching

1. Introduction

Lexical competence is one of the linguistic aspects of communicative competence. It is "the ability to recognize and use words in a language in the way that speakers of the language use them" (Orwig, 1999, p.1). According to Common European Framework of Reference (2001), lexical competence consists of lexical and grammatical elements. Lexical elements include fixed expressions (sentential formulae, phrasal idioms, fixed frames, other fixed phrases such as compound prepositions, fixed collocations) and single word forms (members of the open word classes and closed lexical sets such as days of the week, etc.). Closed word classes (articles, quantifiers, demonstratives, personal pronouns, etc.) belong to grammatical elements.

Some applied linguists have tried to define the different aspects of lexical competence and vocabulary knowledge. Richards (1976) discussed linguistic, psycholinguistic and sociolinguistic aspects of word knowledge. Nation (1990) identified eight types of word knowledge (e.g. form, position, function and relation with other words). According to Qian (2002), vocabulary knowledge should contain four dimensions (vocabulary size, depth of vocabulary knowledge, lexical organization, and automaticity of receptive-productive knowledge). Thus, there is a common consensus that vocabulary knowledge includes at least two dimensions: the size (breadth) and depth of vocabulary knowledge (as cited in Shen, 2008, p. 136) and recent studies usually focus on these two aspects.

In foreign/second language pedagogy, vocabulary teaching has an important place for fluency and accuracy. When compared with traditional approaches and methods, Data Driven Learning (DDL) has been suggested as an effective instructional approach to develop vocabulary knowledge (Chan & Liou, 2005; Ko, 2005). DDL describes the methodology of the use of linguistic corpora in language learning. DDL was seen as a valuable tool to enhance language learning autonomy or independent learning (Sun & Wang, 2003; Thurston & Candlin, 1998). It provides a variety of authentic language data which raise EFL learners' analytical power and independence through self-accessible authentic texts in concordancing program. According to Gabrielatos (2005), it is better to integrate DDL into language classrooms with task-based approaches. Within this frame, the previous studies generally focus on learners' actual performances such as its effect on collocations or synonyms. However, there is limited research data about the effectiveness of task-based DDL instruction in EFL classrooms. Different than the other studies, the present study offers tasks designed specifically for lexical development and intends to investigate Turkish EFL learners' attitudes and beliefs about the

effectiveness of task-based DDL instruction in terms of enhancing awareness and development of lexical knowledge. Therefore, it is hoped that this study would be helpful for language teachers and researchers in giving insights in terms of integrating DDL into current language courses and for corpus software designers in considering the opinions of language learners. The research questions that guided the study were:

1. What are the attitudes of EFL learners towards the effectiveness of DDL tasks on lexical awareness and development?
2. What are the attitudes of EFL learners about positive and negative aspects of corpus use?
3. What are the attitudes and beliefs of EFL learners about task-based DDL instruction?

1.1 Corpora in Language Teaching

It has been acknowledged that the use of corpora can be placed in the mainstream of language learning and teaching due to its pedagogical potential (McEnery, Xiao & Tono, 2006). In terms of the methodology of the use of the corpus examples in language learning and teaching, DDL method was introduced by Johns (1990). In this aspect, pedagogical corpus applications are categorized into two: direct and indirect applications (Leech, 1997; Römer, 2011). Indirect applications of corpus include mainly the researchers and material writers, by focusing on the effects of the corpus on the teaching syllabus and reference works and teaching materials. However, direct applications of corpus cover the teacher-corpus and the learner-corpus interactions.

There has been unquestionably substantial research on the use of DDL for several areas of language teaching; mainly lexis and grammar and their reflection in writing and speaking (Boulton, 2009, 2010; Cotos, 2014; Gabrielatos, 2013; Geluso & Yamaguchi, 2014). DDL was found more advantageous in language teaching when compared to traditional approaches, particularly in improving lexical competence of language learners. Frankenberg-Garcia (2014) reported that learners had the opportunity to access more examples with DDL. When compared to the use of online dictionaries, examples from concordancers have led better acquisition of academic word knowledge and its reflection in writing tasks of language learners (Kaur & Hegelheimer, 2005). Çelik (2011) also found that in terms of retention of collocations, students who used DDL performed better than the ones who used online dictionaries. Furthermore, Chun-guang (2014) found that corpus-driven lexical chunk instruction was helpful to improve writing performance and the use of lexical chunks in writing.

Furthermore, Johns (1990) pointed out that learners' individual experiences with concordance data help them to develop inductive strategies and to become better language learners without being restricted in the classroom. Thus, corpus environment in language teaching may adapt the view of discovery learning (Bernardini, 2004). Furthermore, Gabrielatos (2005) suggested that DDL activities can be integrated into language classrooms with task-based approaches. Within task-based framework, corpus and concordance data can be used in pre- or post-tasks which will increase noticing and awareness towards the language features.

Recent discussion about DDL in language teaching is increasing particularly about how a language learner can benefit from corpus tools and techniques and their attitudes towards these methods (Boulton & Perez-Paredes, 2014). Several studies reported that the outcomes of corpus-based learning in terms of lexical competence and students' attitudes are positive. According to Liu and Jiang (2009) and Yoon and Hirvela (2004), students stated that the corpus techniques helped them improve their writing and awareness of the importance of contextual factors in lexico-grammar. Moreover, Geluso and Yamaguchi (2014) stated that DDL activities are also helpful to improve the use of formulaic language for spoken fluency of the students. On the other hand, some studies reported that DDL has certain drawbacks such as the limitation of application to mostly advanced language learners, technical difficulties of the software, time-consuming nature of DDL; lack of necessary training (Boulton, 2009, 2010; Chambers & O'Sullivan, 2004; Yoon & Hirvela, 2004).

2. Methodology

2.1 Setting and Subjects

There were 126 participants (76.8 % were female; 23.3 % were male) in this study. All of them were upper-intermediate (B2 level as for CEFR) level juniors studying in the English Language Teaching Department (ELT) at a large state university in Turkey. The study took place in the Lexical Competence course as part of their syllabus in the academic year 2013/2014. This was a semester-long, a first grade compulsory course of ELT departments in Turkey. The course took place once a week for 14 weeks, with each session lasting 150 minutes. All participants in this study volunteered for the questionnaire and the focus-group interview. Written consent forms were taken from each participant.

2.2 Design

The content of the Lexical Competence course included vocabulary learning strategies, affixes, synonyms and antonyms, collocations, denotation and connotation and idioms. Each week, the topics were presented to the students in class by the instructors, who were also the researchers. The resources for the course were a course pack developed and gathered

by the instructors regarding the issue of lexical competence and the book entitled “Teaching Vocabulary through Data-Driven Learning” by Shaw (2011) and video tutorials on YouTube. The Corpus of Contemporary American English (hereafter COCA) (Davies, 2008) was chosen for DDL practices since COCA is the largest freely-available contemporary corpus of English.

After first three weeks of the course, the use of corpus was introduced to the students. DDL trainings in the class were conducted along with the peculiar lexical competence topics such as affixes, collocations, synonyms and idioms. At the beginning of the fourth week, the instructor spent up to 50 minutes per week about the use of corpus for the relevant lexical topic. DDL trainings took approximately 10 hours. The instructors acted as facilitators to introduce the use of corpus and to raise their awareness about corpus use. The students were given certain DDL tasks (see Appendix) about the relevant topics after they were introduced to corpus so that DDL tasks were integrated into the course syllabus as part of the course requirement. Within the tasks, the students were mostly given the choices to search the lexical items that they needed or wanted to search or learn more outside the class. Each week, classroom discussions were held about DDL tasks to be accomplished so as to provide students to share their experiences with their peers.

2.3 Data Collection Instruments

Mixed methods design was used as quantitative and qualitative methods lead to data triangulation of the findings. As for the quantitative data, participants were asked to respond to a questionnaire to indicate their perceptions about DDL tasks, and then a focus-group interview was conducted for qualitative data.

2.3.1 Questionnaire

The questionnaire consisted of 43 items of which were rated on a 5-point Likert scale (1=strongly disagree, 5=strongly agree). Some of the items in the questionnaire were adapted from three different studies about using corpora for writing and speaking (Geluso & Yamaguchi, 2014; Liu & Jiang, 2009; Yoon & Hirvela, 2004). The majority of the items were designed by the researchers according to the related literature and the nature of the study. In order to prevent any misunderstanding and to increase the reliability of the study, the questionnaire was presented in participants' native language (Turkish). Before the main study, the questionnaire was piloted with 30 students who were assumed to be the same as the target group to estimate the reliability and internal consistency of the items. Reliability coefficients (Cronbach Alpha) of the scale were .920 and the lowest value for the item-total correlation was 0.43 which indicated the high reliability of the scale.

The Kaiser-Meyer-Olkin (KMO) test and Bartlett's sphericity test were also used to assess whether the factor analysis was appropriate to test the scale structure. KMO test value was found .915 and Bartlett test significance value was ($p = 0.000 / \chi^2=4678.3542; p<0.01$) which indicated that the correlation between the items were meaningful. The 43 items explained 65.78 % of total variance. According to the factor analysis, the questionnaire was divided into five categories: Learners' attitudes towards the effectiveness of DDL tasks in terms of lexical awareness and development; positive aspects of using corpus; the difficulties confronted while doing the tasks through DDL; the effectiveness and delivery of task-based DDL instruction and beliefs about DDL instruction.

2.3.2 Focus Group Interview

At the end of the course, 15 students randomly selected from the participants were invited to participate in a focus group interview. The researchers moderated the focus group interview which lasted 30 minutes. Examples of interview questions included: “What do you think about DDL tasks in terms of lexical awareness and lexical development?”, “Did you find it difficult or easy to use corpus? What kind of difficulties did you face while using corpus to do the tasks?”, “Do you think that using corpus was effective?” and “What do you think about COCA program? The focus group interview was carried out at the university, video recorded, and transcribed by the researchers.

3. Data Analysis

3.1 Questionnaire

The present study focuses on the attitudes of the Turkish EFL learners towards the use of DDL tasks for raising lexical awareness. Means and standard deviations were calculated for the analysis of participants' response to each statement. Factor analysis and correlation analysis were conducted to find out the categories of the questionnaire and the relationships among them. Additionally t-test was applied to find out the differences between the responses of participants for each statement. The median (interquartile range) and arithmetic mean (standard deviation) of the prefitting scores of the five-likert type of questionnaire were $4/5=0,80$.

3.2 Focus Group Interview

The data collected from the focus-group interview were analyzed qualitatively, particularly through content analysis which is a type of analysis including coding for themes, searching patterns, and making interpretations to draw

conclusions (Ellis & Barkhuizen, 2005). The themes were defined by the three researchers to determine the reliability of the qualitative analysis.

4. Results

4.1 Questionnaire

Table 1. The categories of the questionnaire

	The statements assessing learners'	mean	SD	t	p
1	attitudes towards the effectiveness of DDL tasks in terms of lexical awareness and development	3.35	.98540	38.03	.000
2	attitudes towards positive aspects of using corpus	3.23	.77310	46.72	.000
3	attitudes towards the difficulties confronted while using corpus	3.10	.46907	73.96	.000
4	attitudes towards the delivery of task-based DDL instruction	2.98	.48202	69.27	.000
5	beliefs about task-based DDL instruction	3.14	.72728	48.33	.000

The statements in the questionnaire were grouped under five categories as shown in Table 1.

RQ1. What are the attitudes of EFL learners towards the effectiveness of DDL tasks on lexical awareness and development?

Regarding the first category which assesses the self-reported effects of DDL tasks on learners' lexical awareness and development, the statement *my awareness of synonyms in different registers has increased* was the most favored one (m : 3.57).

Table 2. Attitudes towards the effectiveness of DDL tasks on lexical awareness and development

Items	Mean	SD
Through DDL tasks my awareness about		
1. the use of words in different registers has developed.	3.46	1.161
3. the use of idioms in different registers has developed.	3.45	1.139
8. the use of collocations in different registers has developed	3.50	1.161
12. the use of synonyms in different registers has developed.	3.57	1.124
37. the importance of word frequency has developed.	2.95	1.319
Through DDL tasks my overall knowledge of		
5. English has developed.	3.24	1.214
20. synonyms has increased.	3.54	1.118
25. collocations has increased.	3.37	1.168
27. words has increased.	3.26	1.179
29. idioms has increased.	3.23	1.271
42. affixes has increased.	3.45	1.228
6. vocabulary learning strategies has developed.	3.24	1.174

As shown in Table 2, *my knowledge of synonyms has increased* was also a highly accepted statement (m : 3.53) by the participants. The second highly favored statement was *my awareness of collocations in different registers has increased* (m : 3.5). Among the other lexical elements *idioms* have been reported as the least improved one in terms of developing an awareness and knowledge (m : 3.45; m : 3.23). Interestingly, the learners with the moderate rate (m : 3.24) informed that DDL tasks helped them to develop their own vocabulary learning strategies. Additionally, only a limited number of students (m : 2.95) reported that their *awareness towards the importance of word frequency* was developed (Table 2).

RQ2. What are the attitudes of EFL learners about positive and negative aspects of corpus use?

Table 3. Positive aspects of using corpus

Items	Mean	SD
2. Useful for learning vocabulary.	3.32	1.229
11. Useful for language learning.	3.46	1.215
14. Finding collocations easily.	2.95	1.165
28. Finding idioms easily.	1.58	1.297
33. Useful for finding different usages of the same word.	3.34	1.184
35. Finding affixes easily.	3.25	1.203
39. Finding synonyms easily.	2.90	1.343

Regarding the positive and negative aspects of corpus use, as shown in Table 3, it was indicated that the students found corpus *useful for language learning* (m : 3.46) which was one of the highly agreed positive aspects. They also found corpus *useful for finding different usages of the same word* (m : 3.34). The thirdly positive aspect stated by the participants was *corpus is useful for learning vocabulary* (m : 3.32). On the other hand, *finding idioms easily* (m : 1.58) was the least agreed statement among the others.

Table 4. Difficulties confronted while using corpus

Items	Mean	SD
7. Understanding the user guide of COCA.	3.65	1.207
16. Time consuming to search for words.	3.11	1.296
13. Using wildcards.	3.31	1.234
17. Too many sentences.	3.93	.960
18. Too many irrelevant lexical items around key words.	2.73	1.207
21. Access to COCA web page.	2.70	1.270
22. Understanding the concordance lines.	2.71	1.113
24. Learning to use COCA.	3.10	1.227

As for the negative aspects of the corpus use, the participants generally focused on the difficulties confronted while using corpus, as shown in Table 4. The students stated that they had the most difficulty in *searching words among several texts* (m : 3.93), which is followed by *understanding the user guide of COCA* (m : 3.65). They also agreed on the statement about the difficulty of *using wildcards* while using corpus (m : 3.31).

RQ3. What are the attitudes and beliefs of EFL learners about task-based DDL instruction?

Table 5. The attitudes towards the delivery of task-based DDL instruction

Items	Mean	SD
9. It was not beneficial for me to have task-based DDL instruction.	2.88	1.423
15. How to do the tasks by using corpus was sufficiently explained.	3.43	1.073
26. I had difficulty in doing tasks through DDL.	3.49	1.154
30. The information given about the use of DDL was sufficient.	3.22	1.128
31. DDL tasks were not explained sufficiently.	2.45	1.081
32. The in-class activities about the use of DDL were not sufficient.	2.48	1.029

As shown in Table 5, the students pointed out that they had *difficulty in doing tasks through DDL* (m : 3.49). They also agreed on the statement *how to do tasks by using corpus was sufficiently explained* with the mean (3.43). Furthermore, the least favored statement was *DDL tasks were not explained sufficiently*. In general the participants' attitudes were positive about the delivery of tasks through DDL.

Table 6. Students' beliefs about using corpus

Items	Mean	SD
4. I believe that I can easily use corpus when necessary.	3.41	1.185
10. I think it would be better if the use of corpus is supported with a dictionary.	3.68	1.082
19. I didn't like searching for words in corpus.	3.22	1.307
23. I liked DDL instruction.	2.94	1.275
34. I think corpus should also be used in teaching other skills.	2.70	1.276
36. I think corpus is better than printed dictionary.	3.03	1.356
38. I think corpus is better than on-line dictionary.	3.34	1.345
40. I will use corpus whenever I need.	3.38	1.293
41. Using corpus has motivated me to learn words.	2.89	1.200
43. I would suggest DDL to other English learners.	2.85	1.308

Moreover Table 6 reveals learners' beliefs about DDL instruction and corpus use. The mostly agreed statement was *it would be better if the use of corpus is supported with a dictionary* (m : 3.68). They also believed that *they can easily use corpus when necessary* (m : 3.41). However, only a limited number of the participants believed that *corpus should also be used in teaching other skills (reading, writing etc.)* (m : 2.70).

4.2 Focus Group Interview

The data gathered with the focus-group interview were categorized under the following themes:

Using DDL tasks to raise lexical awareness and development

The majority of the students found DDL tasks effective and necessary. They mentioned that they learnt different usages of the words in different contexts. Instead of learning vocabulary in isolated forms, they stated the benefits of learning vocabulary in context. Among the tasks assigned, tasks about synonyms and collocations were the mostly favoured one while the least favoured one was Task 5 about idioms. As for suggestions, they stated that DDL should be integrated into in-class activities together with the teacher. Below is given some quotations about these tasks:

Especially making a word web was so effective for me. While doing the word web, I enjoyed finding the synonyms of the words I have never seen before. (S5)

I think it was very useful for searching collocations in DDL. It was very easy to remember the collocations I have learnt in contexts. I visualized the sentences. Before this, collocations were very challenging for me but now I can remember

them easily. (S7)

The difficulties of using corpus

All of the students reflected that COCA was not user-friendly by listing the following difficulties: wildcards were complicated; help section was not very informative and sufficient; the interface of COCA had frequent technical problems and there was not a proper manual; it was lack of audio recordings of the texts; it seems a kind of program for the professionals and native speakers rather than for the foreign language learners. Moreover, they suggested that there should be a very informative and comprehensive manual for using COCA; the wildcards should be more representative and simplified to handle.

The majority of the students mentioned that corpus would be more useful if it was used with an online dictionary. Since they sometimes had difficulties in understanding the meanings of the words or phrases in corpus, they needed to consult to an online dictionary. Besides, few students reported that using corpus was time consuming. Below is given some quotations about the difficulties confronted:

Searching for words by using corpus takes time, and also very low daily limit is given to the users in COCA. (S6)

If our teacher hadn't taught us how to use corpus, to be honest, I wouldn't have understood by just reading the manual.

The wildcards were also so difficult to keep in mind. (S3)

In my opinion it takes a long time to search for words using corpus. (S4).

Guessing the meanings of the words might be sometimes difficult. I think corpus must be supported by an online dictionary. (S2)

Using corpus for several other purposes

Most of the students reported that corpus use could be beneficial for translation studies. Particularly, it would help them to use the close synonyms in the appropriate context. Some of the students stated that they could use corpus for essay writing and grammar learning. Moreover, some of the students mentioned their preferences of being guided by the teacher. They indicated that they preferred doing the tasks in the classroom with the help of the teacher. They interestingly underlined that they were accustomed to be directed by the teacher in doing such tasks. Even the lexical items to be learnt had been dictated by the teacher without their own preference; thus they had difficulty in controlling their own learning process.

I wish we could have done the tasks in the classroom with the teacher guidance. (S2)

Using the corpus in the class with the help of the teacher was more enjoyable. (S4)

5. Discussion

According to the results, the students agreed that their awareness towards lexical items, particularly their depth of vocabulary was increased. In particular, among DDL tasks, the most effective ones were about synonym searching. Thus, the students reported that DDL was effective for learning synonyms. This is consistent with Kayaoğlu (2013) who found out that the use of corpus for deciding on close synonyms proved to be very effective on participants' performance on the vocabulary portion of the pre-test and post-test. Similarly, corpus use was reported to be effective on the development of learners' knowledge about collocations and affixes. The results also indicated that their awareness towards the use of collocations in different types of texts raised moderately. However, their awareness about the issues such as word frequency, idioms and vocabulary learning strategies did not increase satisfactorily.

Furthermore, the students had positive perceptions about the effectiveness of using corpus for language learning while they did not have strong agreement over the issues such as deficiencies, and comprehensiveness of corpus and its time-consuming nature. These results were consistent with Chambers (2005) who concluded overall positive reactions along with some aspects of disadvantages of using corpus. Regarding the corpus used in the study, COCA, the students reported some difficulties as concluded also by Geluso and Yamaguchi (2014). The results indicated that COCA was not found user-friendly since it did not provide an effective user guide and easy-to-follow wildcards. Thus, as Braun (2005) suggests, there is need for pedagogically relevant corpora by enriching the corpora for more direct uses of learners.

Similar with the previous studies (Sun, 2000; Thurstun & Candlin, 1998), messiness of the data may cause difficulties in interpreting several concordance lines of KWICs. These difficulties could not be attributed to the proficiency level of learners who were upper-intermediate because they reported that they understood texts to a certain level but the presence of several texts for a single word was frustrating for them. This is in line with Farr's (2008) and Kennedy and Miceli's (2001) studies, as learners stated difficulties in refining their searches and interpreting the results. Some of the students found particular tasks (Task 1 about word frequency and Task 5 about idioms) difficult and particularly in the focus group interview they suggested that corpora should be used in-class activities together with the teacher. This can be explained in two ways. First, they might find it demanding to search for substantial amounts of targeted practice on selected items and to refine their queries with subsequent searches (Kennedy & Miceli, 2001) since they were generally

accustomed to be provided with invented and impoverished texts. Second, the nature of the tasks chosen for the study was not clear-cut which would limit their discovery process. However, EFL learners who have limited exposure to target language were familiar with traditional in-class activities or distinctly and sharply defined teacher-lead tasks. Though more learner-centred tasks which would promote self-discovery were chosen by the researchers, some of the learners still found these tasks time-consuming and reported limited benefit of corpus use particularly on developing their own vocabulary learning strategies.

Furthermore, the findings showed that the students supported the idea that corpus use and dictionary use should be integrated. This is consistent with Yoon and Hirvela (2004) who found the students' suggestions about the complementary aspects of both resources. Thus, while dictionaries are time-effective for searching words, the corpus provides more contextual evidences about a particular word with its several instances. However, Kaur and Hegelheimer (2005) and Çelik (2011) found that corpus use led better acquisition of lexical items when compared with online dictionary uses.

Regarding the delivery of DDL instruction, the results indicated that the students found the integration of DDL instruction into the course and its training adequate. It can be concluded that corpus-consulting skills could be improved by familiarization with DDL. In this sense, the findings are in line with Frankenberg-Garcia (2012) particularly about the necessity of efficient training about corpus use.

6. Conclusion and Implications

It has been acknowledged that lexical awareness covers different aspects of lexical competence and vocabulary knowledge such as vocabulary size, depth of vocabulary knowledge, lexical organization, and automaticity of receptive-productive knowledge. However, the findings of the current study revealed that the learners' lexical awareness raised only for certain aspects related to receptive depth of vocabulary knowledge. Though it was not the scope of this study to assess actual performances of the learners about various aspects of lexical competence we can draw some important conclusions for further studies which will focus on developing lexical competence.

In EFL pedagogy, corpora have been used both directly and indirectly (Römer, 2011). Rather than consulting to corpora indirectly such as developing curriculum and materials and obtaining authentic material, this study made use of DDL directly. Since it has been suggested by several scholars that corpus consultation might be more beneficial when used outside the classroom as it offers students a high degree of autonomy (Chambers, 2007; Yoon & Hirvela, 2004), tasks involved in this study were designed not only to improve learners' lexical competence but also their autonomy. However, unlike previous studies, most of the students showed resistance to DDL tasks designed with the aim of increasing their awareness about different aspects of lexical competence and vocabulary knowledge. They insisted on using DDL tasks in the classroom under the control of the teacher. Thus, for more autonomous learners in terms of corpus use, guidance is needed on how to use corpus with exploratory tasks and opportunities to experience corpus skills (Braun, 2005).

Promoting the use of DDL into EFL education mostly depends on teachers (Farr, 2008; Tsui, 2005), materials developers and corpus designers. Teachers should be aware of the fact that time spent on DDL is not time wasted. Furthermore, corpus designers should not underestimate the complaints of the users depending on the software which was found complicated and not user-friendly. The manual and the introduction of corpora with integrated interfaces for on-line access can be re-designed by considering non-native speakers as well. As suggested by several scholars such as Chujo, Utiyama and Nishigaki (2007) and Wible et al. (2002), the messiness of the texts may be avoided by grading the texts within a corpus or the concordance output with a more accessible language if it is possible. In any case, an awareness would increase about the benefits of corpora in EFL education not as substituting the existing methods and approaches, but as a favourable support.

References

- Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. Sinclair (Ed) *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins, 15-36. <http://dx.doi.org/10.1075/scl.12.05ber>
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(1), 37–54. <http://dx.doi.org/10.1017/S0958344009000068>
- Boulton, A. (2010). Data-Driven Learning: Taking the computer out of the equation. *Language Learning*, 60(3), 534-572. <http://dx.doi.org/10.1111/j.1467-9922.2010.00566.x>
- Boulton, A., & Pérez-Paredes, P. (2014). Researching uses of corpora for language teaching and learning. *ReCALL*, 26(2), 121-127. <http://dx.doi.org/10.1017/S0958344014000068>

- Braun, S. (2005). From pedagogically relevant corpora to authentic language learning contents. *ReCALL*, 17(1), 47-64. <http://dx.doi.org/10.1017/S0958344005000510>
- Çelik, S. (2011). Developing collocational competence through web based concordance activities. *Novitas-ROYAL (Research on Youth and Language)*, 5(2), 273-286.
- Chambers, A. (2005). Integrating Corpus Consultation in Language Studies. *Language Learning and Technology*, 9(2), 111-125.
- Chambers, A. (2007). Popularising corpus consultation by language learners and teachers. In E. Hidalgo, L. Quereda, J. Santana (Eds) *Corpora in the Foreign Language Classroom: Selected Papers from the Sixth International Conference on Teaching and Language Corpora (TaLC 6)* (pp. 3-16) Amsterdam: Rodopi. http://dx.doi.org/10.1163/9789401203906_002
- Chambers, A., & O'Sullivan, Í. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL*, 16(1), 158-172. <http://dx.doi.org/10.1017/S0958344004001211>
- Chan, T., & Liou, H. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer-Assisted Language Learning*, 18(3), 231-250. <http://dx.doi.org/10.1080/09588220500185769>
- Chujo, K., Utiyama, M., & Nishigaki, C. (2007). Towards building a usable corpus collection for the ELT classroom. In E. Hidalgo, L. Quereda, J. Santana (Eds) *Corpora in the Foreign Language Classroom* (pp. 47-69). Amsterdam: Rodopi. http://dx.doi.org/10.1163/9789401203906_005
- Chun-guang, T. (2014). An Empirical Research on the corpus-driven lexical chunks instruction. *International Journal of English Language Teaching*, 2(2), 1-36.
- Cotos, E. (2014). Enhancing writing pedagogy with learner corpus data. *ReCALL*, 26(2), 202-224. <http://dx.doi.org/10.1017/S0958344014000019>
- Council of Europe. (2001). *The Common European framework of reference for languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Davies, M. (2008). *The Corpus of Contemporary American English: 450 million words, 1990-present*, available online at <http://corpus.byu.edu/coca/>
- Ellis, R., & Barkhuizen, G. (2005). *Analysing Learner Language*. Oxford: Oxford University Press.
- Farr, F. (2008). Evaluating the use of corpus-based instruction in a language teacher education context: Perspectives from the users. *Language Awareness*, 17(1), 25-43. <http://dx.doi.org/10.2167/la414.0>
- Frankenberg-Garcia, A. (2012). Learners' use of corpus examples. *International Journal of Lexicography*, 25(3), 273-296. <http://dx.doi.org/10.1093/ijl/ecs011>
- Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL*, 26(2), 128-146. <http://dx.doi.org/10.1017/S0958344014000093>
- Gabrielatos, C. (2005). Corpora and language teaching: Just a fling, or wedding bells? *TESL-EJ*, 8(4), 1-37.
- Gabrielatos, C. (2013). If-conditionals in ICLE and the BNC: A success story for teaching or learning? In S. Granger, G. Gilquin, F. Meunier (Eds.) *Twenty Years of Learner Corpus Research: Looking back, moving ahead* (pp. 155-166). Belgium: Presses Universitaires de Louvain.
- Geluso, J., & Yamaguchia, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, 26(2), 225-242. <http://dx.doi.org/10.1017/S0958344014000044>
- Johns, T. (1990). From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. *CALL Austria*, 10, 14-34.
- Kaur, J., & Hegelheimer, V. (2005). ESL students' use of concordance in the transfer of academic word knowledge: An exploratory study. *Computer Assisted Language Learning*, 18(4), 287-310. <http://dx.doi.org/10.1080/09588220500280412>
- Kayaoğlu, M. N. (2013). The use of corpus for close synonyms. *The Journal of Language and Linguistic Studies*, 9(1), 128-144.
- Kennedy, C., & Miceli, T. (2001). An evaluation of intermediate students' approaches to corpus investigation. *Language Learning and Technology*, 5(3), 77-90.
- Ko, M. H. (2005). Korean university student use of phrasal and one-word verbs and concordance-based learning.

Korean Journal of Applied Linguistics, 21(2), 1-39.

- Leech, G. (1997). Teaching and language corpora: a convergence. A. Wichmann, S. Fligelstone, T. McEnery & G. Knowles (Eds.) *Teaching and Language Corpora* (pp. 1-23). London: Longman.
- Liu, D., & Jiang, P. (2009). Using a corpus-based lexicogrammatical approach to grammar instruction in EFL and ESL contexts. *The Modern Language Journal*, 93(1), 61-78. <http://dx.doi.org/10.1111/j.1540-4781.2009.00828.x>
- McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-Based Language Studies: An Advanced Resource Book*. London: Routledge.
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. New York: NY Newbury House.
- Orwig, C. (1999). *Aspects of Communicative Competence*. Extract from the Lingualinks library, version 3.5, published on CD-Rom by SIL international.
- Qian, D. D. (2002). Investigating the relationship between the vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513-536. <http://dx.doi.org/10.1111/1467-9922.00193>
- Richards, J. C. (1976). The role of vocabulary teaching. *TESOL Quarterly*, 10(1), 77-90. <http://dx.doi.org/10.2307/3585941>
- Römer, U. (2011). Corpus research applications in second language teaching. *Annual Review of Applied Linguistics*, 31, 205-225. <http://dx.doi.org/10.1017/S0267190511000055>
- Shaw, E. M. (2011). *Teaching Vocabulary through Data-Driven Learning*. Brigham Young University, USA.
- Shen, Z. (2008). The roles of depth and breadth of vocabulary knowledge in EFL Reading Performance. *Asian Social Science*, 4(12), 135-137.
- Sun, Y. C. (2000). Using on-line corpus to facilitate language learning. Paper presented at *the Annual Meeting of the Teachers of English to Speakers of Other Languages*, British Columbia, Canada.
- Sun, Y., & Wang, L. (2003). Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer-Assisted Language Learning*, 16(1), 83-94. <http://dx.doi.org/10.1076/call.16.1.83.15528>
- Thurstun, J., & Candlin, C. (1998). *Concordancing and the teaching of the vocabulary of academic English*. *English for Specific Purposes*, 17(3), 267-280. [http://dx.doi.org/10.1016/S0889-4906\(97\)00013-6](http://dx.doi.org/10.1016/S0889-4906(97)00013-6)
- Tsui, A. (2005). ESL teachers' questions and corpus evidence. *International Journal of Corpus Linguistics*, 10(3), 335-356. <http://dx.doi.org/10.1075/ijcl.10.3.03tsu>
- Wible, D., Chien, F., Kuo, C. H., & Wang, C. (2002). Toward automating a personalized concordancer for data-driven learning: A lexical difficulty filter for language learners. In B. Kettemann, G. Marko (Eds) *Teaching and Learning by Doing Corpus Analysis* (pp. 147-154). Amsterdam: Rodopi.
- Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing*, 13(4), 257-283. <http://dx.doi.org/10.1016/j.jslw.2004.06.002>

APPENDIX: Design of DDL Tasks

Objectives		
To enable students to learn lexical items in context, To help students study more effectively on their own. To expand the breadth and depth of vocabulary knowledge of the students, To raise awareness about vocabulary learning, To increase awareness of the students towards the register of the lexical items, To raise students' awareness about easy access of searching for lexical items.		
Explanation		Aim
Task 1	* Search for at least five words about their frequency level in different registers of written and spoken texts (academic, fiction, magazine and newspaper).	* To raise students' metalinguistic awareness of frequency and register embedded in a specific context in vocabulary learning. * To help them to decide in choosing the most useful words for the appropriate registers.
Task 2	* Select five prefixes and five suffixes which you want to search more chosen from the in-class activities and search them in corpus. * Analyse the findings in terms of their roots and affixes with their meanings. * Write reports about the findings including the example sentences which make the meaning clearer for you.	* To raise students' metalinguistic awareness of morphology through corpus. * To make students notice the meanings of the affixes and their effects on the words.
Task 3	* Select ten words from either in-class activities or any other words you want to search for. * By using COCA, find out the synonyms and example sentences for each word. * After choosing only one of the words from your list, organize a word web of the selected word via COCA.	* To help students access to more information in corpus about the synonymous words which is limitedly provided in thesaurus. * To help students develop metalinguistic knowledge such as register, part of speech, and knowledge of collocations and connotation, which leads to synonym knowledge via corpus.
Task 4	* Search the collocations of 10 frequent academic words (<i>indicate, process, required, research, significant, specific, major, issues, role, principle</i>). * Illustrate the findings in a table consisting the definition, collocates and sample phrases in example sentences.	* To increase students' collocational awareness via corpus. * To make the students aware of easy access to collocations and their usages via corpus which is limitedly provided in other sources. * To develop students' collocational knowledge about the target academic words.
Task 5	* Select ten idioms which you want to search more taken from the in-class activities and search them in corpus. * Analyse the example sentences in terms of their meanings and registers. * Write reports about the findings including the example sentences which make the meaning clearer for you.	* To make the figurative meanings of idioms more comprehensible for students via corpus. * To make the students aware of easy access to contexts of idioms and their usages via corpus which is limitedly provided in other sources.

